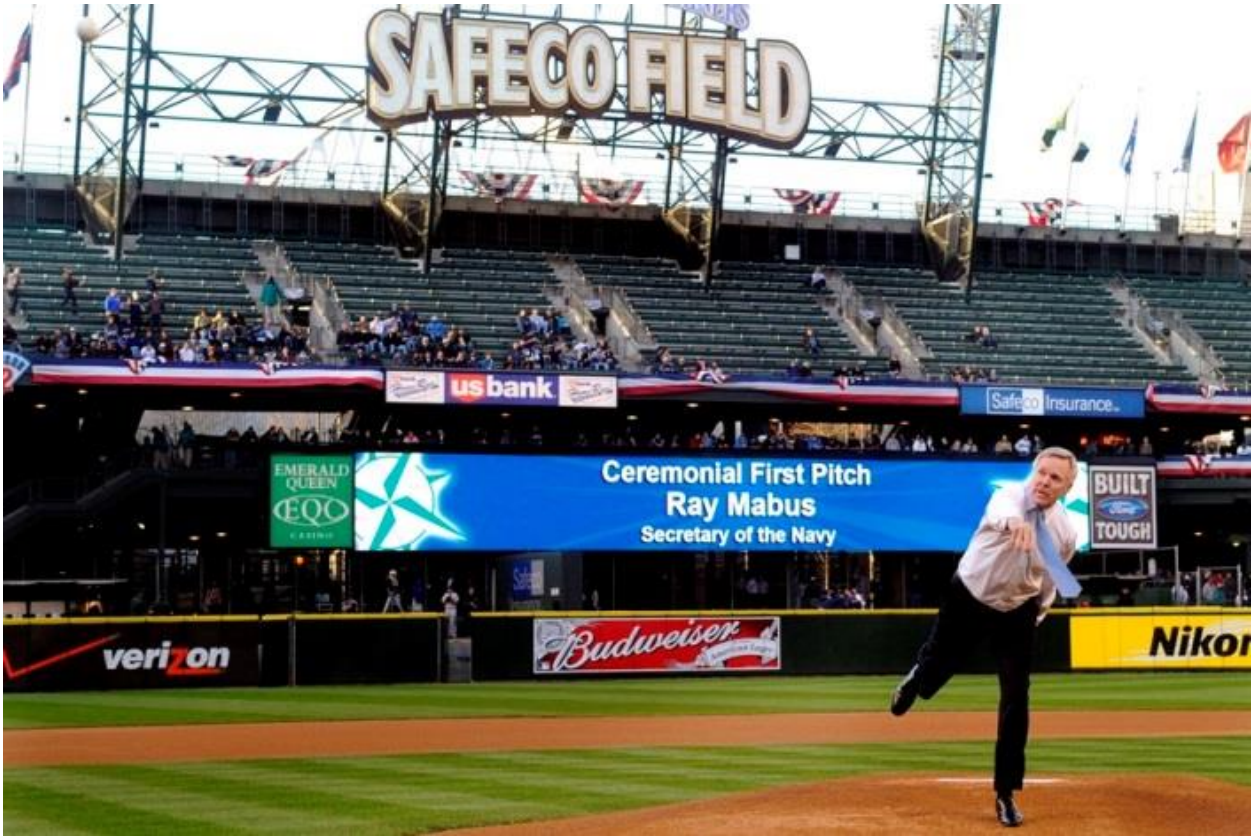
Can “Sabermetrics” Replace the Officer FITREP?



By Robert Kozloski and Major Armando Martinez **

Major League Baseball is a data-rich environment providing a great example of how analytics can effectively assess and manage talent. As outlined in the book (later movie) Moneyball, the traditional role of the human talent scout, a so-called subject matter expert using highly personal judgment, is being replaced by a statistician who understands baseball. Furthermore, the increasing popularity of fantasy sports provides clear evidence of how analytics are permeating American society, a trend arising in many fields except military personnel systems.

Talent decisions in professional baseball were historically viewed through the eyes of an experienced observer, supported by standard scales of performance: Runs, Hits, and Runs-Batted-In (for hitters) or Wins, Strikeouts, and Earned Run Average (for pitchers). Because of the ease of data collection and analysis, more insightful measures of performance were created to make more informed personnel decisions.

Sabermetrics, named in honor of the Society for American Baseball Research, uses statistical analyses to evaluate, compare and forecast baseball talent. Early use of these non-traditional metrics started in the 1970s, but only became popular this past decade, largely due to advances in information sharing using better quantitative techniques. With more team owners—both real and fantasy—relying on analytics, the value of data-driven insight is readily apparent.

One example of how baseball statistics have evolved can be found in the new "Wins Above Replacement" (WAR) measurement. This method compares an individual player to a "replacement player," who could be hired easily and (likely) at the league-minimum salary. WAR takes existing metrics and uses them in a complex algorithm to calculate the number of wins an individual player will bring a team, beyond the replacement's value, by considering contributions on both offense and defense.

Sabermetrics brings objectivity to baseball management through a more robust statistical analysis. Managers can now make personnel decisions not only using the quality of a player, but by assessing how well one fits within the existing talent structure of a team. A team which plays in a hitter-friendly ballpark will target players differently than a team at home in a pitcher-friendly ballpark. A manager is now also more likely to consider the talent surrounding a player—their context within the team—than base a decision solely on the individual player's skills independently.

Like Major League Baseball, the success of the US military depends upon best using the talent of our outstanding men and women in uniform. No baseball team would stand a chance on the field today if it didn't incorporate advanced analytics into the clubhouse; the US military must recognize its shortfall in these tools and put data analytics into its game.

Officer Quantitative Rating Score

How the US military measures and uses officer talent is fundamentally broken, particularly in the Navy, where the majority of officers have lost faith in the current evaluation system. One of the main flaws with our antiquated approach is the over-emphasis on subjective assessments and a lack of clear measures. Such shortfalls could cripple the services, whose "talent scouts" (commanding officers) promote junior officer much like themselves. If the combat environment shifts too quickly, this approach will yield disaster. The naval services, therefore, can learn a great deal from Major League Baseball to resolve this problem.

The first step to creating a data-driven performance system is to define the characteristics desired from naval officers and then develop ways to measure performance. To illustrate the point, the following analysis will use physical fitness, cognitive skills, tactical proficiency and leadership as the variables to calculate the Officer Quantitative Rating Score (OQRS).

Physical Fitness: There are several ways to measure physical fitness. It can easily be appraised with a combination of body composition assessments and fitness tests results. There are other factors which could be used in this algorithm, such as age, sex and deployment time, which can help create a better fitness value. Commercial health monitoring technology, such as the ubiquitous Fitbit type monitor, will eventually be included in this measure.

Cognitive Skills: While cognitive skills may be difficult to quantify, there are a number of component factors which could be measured: standardized tests results, academic performance, professional publications, or practical problem solving tests. Once the data sources are defined, these measures could develop a base "cognitive rating" and officers would work towards improving their ratings over the course of a career.

Tactical Proficiency: This component must be customized for each designator, so that tactical proficiency in a warfighting community could be measured clearly at each skill level from novice to master practitioner. Specific training experience, deployment time and warfare qualifications can all be used to calculate this value.

This variable would also be used to choose between specialist or generalist career paths. Some officers have significant operational experience within their warfare community while others gain operational experience from joint duty assignments. Distinguishing between the two could inform selection decisions for a joint command or senior naval warfare community assignment.

Leadership: Measuring this component is likely the most difficult and contentious. In the past most leadership measures focused on personal accomplishments rather than actually leading teams or organizations. One way to gauge leadership is to establish clearly defined goals, measuring team performance. Beginning each reporting year, officers should be given a set of challenging but achievable goals commensurate with rank, billet and mission. The annual assessment would measure the extent to which goals were accomplished. Additionally, the results of 360 degree performance assessments, to assess leadership behavior, as well as command climate surveys, could feed into this calculation.

Further, types of billets or duty assignments could be categorized and factored into this value. In baseball, a pitcher for the Colorado Rockies is likely to have a higher earned run average than one from a pitcher-friendly home ballpark. This difference can be attributed to the higher altitude of the Rockies' stadium in Denver, allowing more home runs. Each community manager knows some billets are more demanding than others, so an officer meeting her personal goals at a deployed or high tempo unit would stand out

more than one in a less demanding role. Unit assessment reports such as the Defense Readiness Reporting System or Status of Resources and Training System could factor in to this value as well.

Once the OQRS score is calculated, it must be placed into meaningful context. The officer receiving the assessment must be provided the chance to note annual achievements or provide explanation why goals were not met. Similarly, the reporting senior should comment on promotion potential, preferably on a grade-inflation proof scale, and provide insight on the OQRS score for that particular year. Limited resources, for example, may have prevented an officer from achieving a specific goal. Ideally, the OQRS would be used consistently over an entire career, similar to a lifetime batting average in baseball.

Rather than using this score simply for promotions, as is our current practice, the OQRS could help both the service member to achieve career goals while aiding the institution in managing the actual talent of the workforce, rather than simply putting "butts in seats".

OQRS scores should be viewable by all officers, to take personnel evaluation ratings out of "the shadows" and make the entire process transparent. The secrecy and back-room deals surrounding the reporting of fitness are holdovers from a bygone era. Officers will know how they stack up against peers and in what areas they need improvement. Top performing officers would welcome such transparency and weaker officers will get the message to improve or get out.

With such data, Detailers, Monitors and command selection boards can make informed decisions about where people "best fit" for a particular position. Senior commanders, similar to a GM in baseball, could request a specific type of officer to fill the immediate organizational needs. For example, if they know a billet demands strong cognitive skills but less leadership, more weight could be assigned to the first variable to identify officers meeting this criterion.

Like professional baseball, the military personnel system is a data rich environment, yet we have failed to take advantage of our existing information and create a truly effective talent management system. Many may argue that expert judgment is the best way to manage our naval officers, simply out of tradition. However, we do not know how to measure the underutilization of officer talent, nor the number of capable people who exit military service each year because the full potential of officers often stays on the bench.

As an institution, we must admit the current system is not working; it can't be used to build a World Series champion in this new environment. Worse, given the current threat environment we cannot afford to have a "rebuilding year" to resolve this issue. Changing our personnel patterns now may allow us to win our next war. We must explore new data-driven options to measure effectively and use the talent of our naval officers. The sabermetrics approach to talent management in baseball provides an excellent starting point.

Major Armando Martinez is an active duty Marine Corps Officer (and fantasy baseball legend**). Mr. Robert Kozloski is a Senior Program Analyst.

** =The opinions expressed here are solely those of the author(s), and do not necessarily reflect those of the Department of the Navy, Department of Defense or the United States government.